Title

Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production.

Journal

PLoS genetics, 7(3)

ISSN

1553-7390

Authors

Chung, Sharon A
Taylor, Kimberly E
Graham, Robert R
et al.

Publication Date

2011-03-01

DOI

10.1371/journal.pgen.1001323

Peer reviewed

OPEN ACCESS Freely available online

PLoS GENETICS

Differential Genetic Associations for Systemic Lupus Erythematosus Based on Anti–dsDNA Autoantibody Production

Sharon A. Chung[1]*, Kimberly E. Taylor[1], Robert R. Graham[2], Joanne Nititham[1], Annette T. Lee[3], Ward A. Ortmann[2], Chaim O. Jacob[4], Marta E. Alarcón-Riquelme[5,6,7], Betty P. Tsao[8], John B. Harley[9], Patrick M. Gaffney[7], Kathy L. Moser[7], SLEGEN[¶], Michelle Petri[10], F. Yesim Demirci[11], M. Ilyas Kamboh[11], Susan Manzi[12], Peter K. Gregersen[3], Carl D. Langefeld[13], Timothy W. Behrens[2], Lindsey A. Criswell[1]*

1 Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, University of California San Francisco, San Francisco, California, United States of America, **2** ITGR Human Genetics, Genentech, South San Francisco, California, United States of America, **3** Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, Manhasset, New York, United States of America, **4** Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America, **5** Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden, **6** Andalusian Center for Genomics and Oncological Research, Pfizer-University of Granada-Junta de Andalucía, Granada, Spain, **7** Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, United States of America, **8** Department of Medicine, University of California Los Angeles, Los Angeles, California, United States of America, **9** US Department of Veterans Affairs Medical Center and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States of America, **10** Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **11** Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **12** Allegheny-Singer Research Institute, Pittsburgh, Pennsylvania, United States of America, **13** Department of Biostatistical Sciences, Wake Forest University Health Sciences, Wake Forest, North Carolina, United States of America

Abstract

Systemic lupus erythematosus (SLE) is a clinically heterogeneous, systemic autoimmune disease characterized by autoantibody formation. Previously published genome-wide association studies (GWAS) have investigated SLE as a single phenotype. Therefore, we conducted a GWAS to identify genetic factors associated with anti–dsDNA autoantibody production, a SLE–related autoantibody with diagnostic and clinical importance. Using two independent datasets, over 400,000 single nucleotide polymorphisms (SNPs) were studied in a total of 1,717 SLE cases and 4,813 healthy controls. Anti–dsDNA autoantibody positive (anti–dsDNA +, n = 811) and anti–dsDNA autoantibody negative (anti–dsDNA –, n = 906) SLE cases were compared to healthy controls and to each other to identify SNPs associated specifically with these SLE subtypes. SNPs in the previously identified SLE susceptibility loci *STAT4*, *IRF5*, *ITGAM*, and the major histocompatibility complex were strongly associated with anti–dsDNA + SLE. Far fewer and weaker associations were observed for anti–dsDNA – SLE. For example, rs7574865 in *STAT4* had an OR for anti–dsDNA + SLE of 1.77 (95% CI 1.57–1.99, p = 2.0E-20) compared to an OR for anti–dsDNA – SLE of 1.26 (95% CI 1.12–1.41, p = 2.4E-04), with $p_{heterogeneity} < 0.0005$. SNPs in the SLE susceptibility loci *BANK1*, *KIAA1542*, and *UBE2L3* showed evidence of association with anti–dsDNA + SLE and were not associated with anti–dsDNA – SLE. In conclusion, we identified differential genetic associations with SLE based on anti–dsDNA autoantibody production. Many previously identified SLE susceptibility loci may confer disease risk through their role in autoantibody production and be more accurately described as autoantibody propensity loci. Lack of strong SNP associations may suggest that other types of genetic variation or non-genetic factors such as environmental exposures have a greater impact on susceptibility to anti–dsDNA – SLE.

Citation: Chung SA, Taylor KE, Graham RR, Nititham J, Lee AT, et al. (2011) Differential Genetic Associations for Systemic Lupus Erythematosus Based on Anti–dsDNA Autoantibody Production. PLoS Genet 7(3): e1001323. doi:10.1371/journal.pgen.1001323

Editor: Paul I. W. de Bakker, Brigham and Women's Hospital, United States of America

Received July 10, 2010; **Accepted** January 31, 2011; **Published** March 3, 2011

Funding: Support for this study was provided by the NIH grants RR024130 (SAC), AR043815 (COJ), RR020143 (MEA-R), AR043814 (BPT), AR062277 (JBH), RR020143 (JBH), AR042460 (JBH), AI024717 (JBH), AI063274 (PMG), AR052125 (PMG), AR043247 (KLM), AR043727 (MP), AR057028 (FYD), HL092397 (MIK), HL088648 (MIK), HL054900 (MIK), HL074165 (MIK), AR002213 (SM), AR046588 (SM), AR012256 (PKG), AI095386 (PKG), AI068759 (PKG), AR044804 (LAC), AR02175 (LAC), and AR052300 (LAC). Additional support was provided by the American College of Rheumatology Physician Scientist Development Award (SAC), Alliance for Lupus Research (SLEGEN, JBH, KLM, COJ), Knut and Alice Wallenberg Foundation (MEA-R), Torsten and Ragnar Soederbergs Foundation (MEA-R), Swedish Research Council (MEA-R), Gustaf V 80th-Year Jubilee (MEA-R), Swedish Rheumatism Association (MEA-R), Mary Kirkland Scholar Awards (JBH, LAC), US Department of Veterans Affairs (JBH), Lupus Foundation of Minnesota (KLM), and the Rosalind Russell Medical Research Center for Arthritis (LAC, SAC). RRG, WAO, and TWB are employees of Genentech. These studies were performed in part at the University of California San Francisco, General Clinical Research Center (Moffitt Hospital), the Johns Hopkins University General Clinical Research Center, and the University of Pittsburgh General Clinical Research Center with funds provided by the U.S. Public Health Service National Center for Research Resources, M01 RR-000079, M01 RR-000052, and M01-RR000056. The Broad Institute Center for Genotyping and Analysis is supported by grant U54 RR020278 from the National Center for Research Resources. The contents of this project are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sharon.chung@ucsf.edu (SAC); lindsey.criswell@ucsf.edu (LAC)

¶ For membership of The International Consortium on the Genetics of Systemic Lupus Erythematosus please see the Acknowledgments.

Author Summary

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease that can involve virtually any organ system. SLE patients produce antibodies that bind to their own cells and proteins (autoantibodies) which can cause irreversible organ damage. One particular SLE–related autoantibody directed at double-stranded DNA (anti–dsDNA) is associated with kidney involvement and more severe disease. Previous genome-wide association studies (GWAS) in SLE have studied SLE itself, not particular SLE manifestations. Therefore, we conducted this GWAS of anti–dsDNA autoantibody production to identify genetic associations with this clinically important autoantibody. We found that many previously identified SLE–associated genes are more strongly associated with anti–dsDNA autoantibody production than SLE itself, and they may be more accurately described as autoantibody propensity genes. No strong genetic associations were observed for SLE patients who do not produce anti–dsDNA autoantibodies, suggesting that other factors may have more influence in developing this type of SLE. Further investigation of these autoantibody propensity genes may lead to greater insight into the causes of autoantibody production and organ damage in SLE.

Introduction

Systemic lupus erythematosus (SLE) is the prototypic systemic autoimmune disease and can affect virtually any organ system. Manifestations of SLE are quite varied and include renal failure, hemolytic anemia, arterial and venous clots, and disfiguring skin rashes. Overall prevalence of SLE in the general population is 1 in 2000 individuals with a predilection for women (female to male ratio of 6-10:1) [1]. Although the prevalence is relatively low, SLE creates tremendous health care and societal costs since affected individuals are typically young and can suffer significant morbidity and early mortality.

The pivotal immunologic disturbance in SLE is the formation of autoantibodies directed at cell membrane and nuclear components. Deposition of immune complexes containing these autoantibodies leads to inflammatory responses and end-organ damage. Autoantibodies directed against native double-stranded DNA (dsDNA) are observed in 40–60% of SLE patients. Anti–dsDNA autoantibodies can be present prior to clinical symptoms of SLE [2], and are implicated in the pathogenesis of lupus nephritis, a major cause of morbidity and mortality in SLE [3,4]. Anti–dsDNA autoantibodies have also been associated with decreased survival [4]. Given its high specificity for SLE, anti–dsDNA autoantibody production is one of the 11 classification criteria for SLE developed by the American College of Rheumatology (ACR) [5,6].

SLE susceptibility is strongly influenced by both genetic and environmental factors. Recent genetic association studies have successfully identified over 20 SLE susceptibility loci [7]. Odds ratios (OR) for these associations have been modest, with most OR <1.3. One potential factor influencing the magnitude of these associations may be the extensive clinical heterogeneity of SLE. Studying more specifically defined SLE manifestations may reveal stronger and novel genetic associations. Therefore, we conducted a genome-wide association study of anti–dsDNA autoantibody production in SLE to identify genetic associations with this clinically relevant autoantibody, and to determine if the genetic associations were different between those SLE subjects that do and do not produce this autoantibody.

Results

For this genome-wide association study (GWAS), we utilized genotyping data from the GWAS of SLE published by Hom et al. [8] as the discovery dataset, and genotyping data from the GWAS of SLE published by The International Consortium on the Genetics of Systemic Lupus Erythematosus (SLEGEN) [9] as the replication dataset. Since both datasets utilized publicly available healthy controls from the same sources, we supplemented the controls in the replication dataset with 1142 healthy controls from the Cancer Genetic Markers of Susceptibility (CGEMS) study (http://cgems.cancer.gov/data/) [10]. After employing data quality measures, including removal of duplicate and related subjects (see Methods and Figure 1), a total of 1717 SLE cases and 4813 healthy controls of European descent were studied. The discovery dataset was comprised of 1278 SLE cases and 3334 healthy controls, while the replication dataset was comprised of 439 SLE cases and 1479 healthy controls. For both datasets, 47% of the SLE cases were anti–dsDNA +. In the joint dataset, 296,509 SNPs were typed in common between the discovery and replication datasets and passed data quality measures (see Methods). An additional 124,809 imputed SNPs (see Methods) passed data quality filters and were included for analysis in the replication and joint datasets. Figure 1 summarizes the autoantibody status and sample sizes of the datasets used in this study, as well as the number of individuals removed for each data quality measure. The clinical characteristics of the subjects in this study, provided in Table 1, are comparable to those in previously reported studies [1,11].

Anti–dsDNA + SLE versus healthy controls

We first compared anti–dsDNA + SLE cases to healthy controls using additive logistic regression models implemented in PLINK [12] (http://pngu.mgh.harvard.edu/purcell/plink/). The discovery and replication datasets were analyzed separately, and then combined into a "joint analysis" for maximal statistical power. All logistic regression models were adjusted for population stratification using principal components analysis. Table S1 presents the genomic control inflation factor (λ_{GC}) for each analysis prior to and after adjustment for population stratification. P-values for association were adjusted for the λ_{GC} observed after accounting for population stratification (see Methods for additional details). The quantile-quantile and Manhattan plots for the joint analysis are displayed in Figure S1.

Table 2 (and Table S2) displays each locus with significant ($p<5E-07$) or suggestive ($p<1E-05$) evidence of association in the joint analysis. Excluding the associations seen with the major histocompatilibity complex (MHC) on chromosome 6p21, 14 statistically significant associations were observed in the joint analysis of genotyped SNPs when none would have been expected under the null hypothesis. The most significant associations were observed in the MHC, with rs1150754 near *TNXB* (OR_{joint} 2.21, 95% CI 1.92–2.53, $p = 6.4E-29$) as the most significantly associated SNP. Outside of the MHC, the most significantly associated SNP was rs7574865 (OR_{joint} 1.77, 95% CI 1.57–1.99, $p = 2.0E-20$) located in *STAT4*. Strong evidence of association was observed with SNPs in/near *IRF5* and *ITGAM*. Association results for these loci met the genome-wide significance threshold in both datasets, and thus are considered replicated findings. These 3 loci were previously shown to be associated with SLE [8,13,14], but only *STAT4* has been previously associated with anti–dsDNA autoantibody production [15,16]. Strong evidence of association was also observed for *BLK* in the joint analysis, but this association did not meet the threshold for genome-wide significance in both

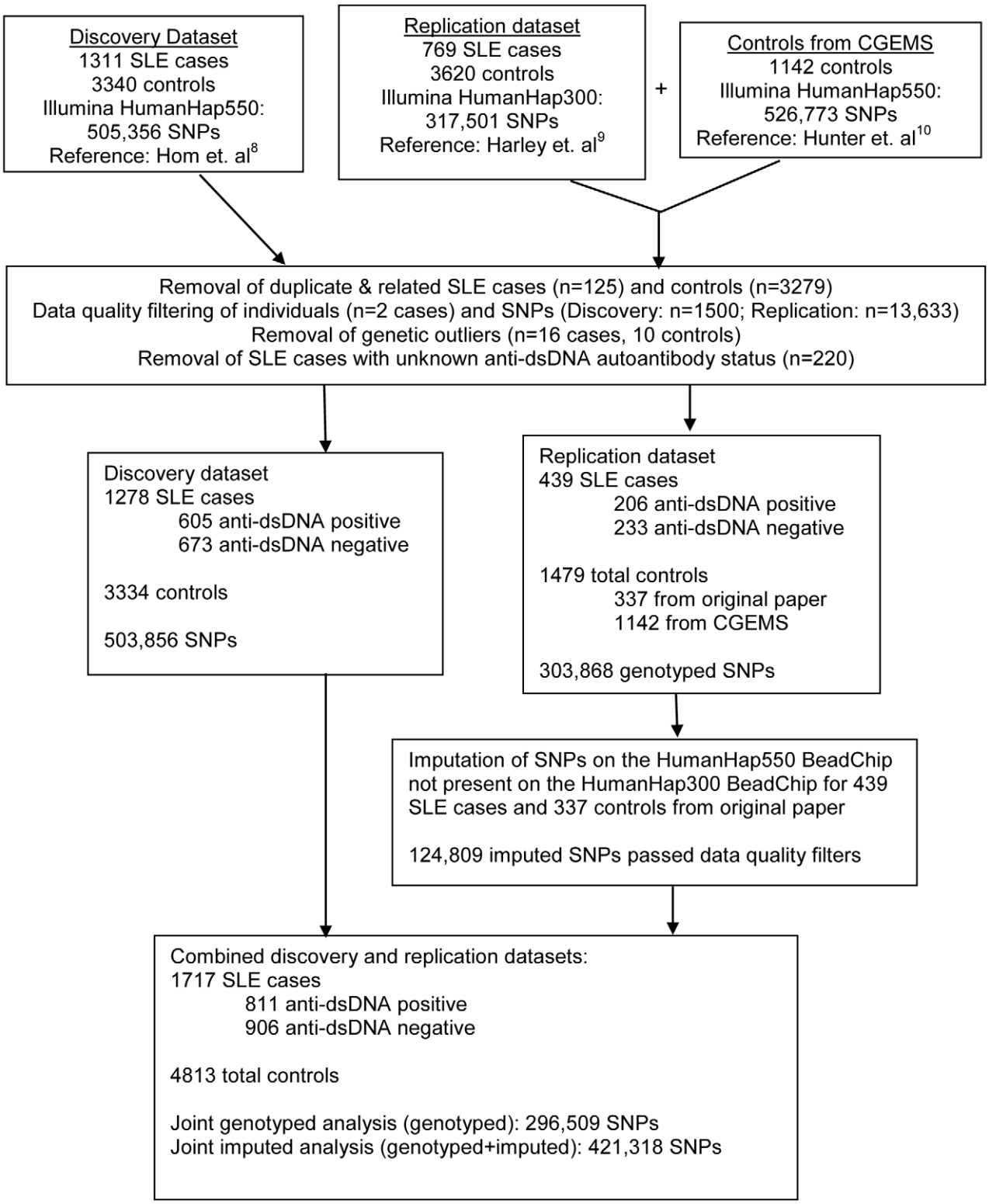

Figure 1. Overview of sample sizes and genotyping platforms used in this study.
doi:10.1371/journal.pgen.1001323.g001

datasets. While SNPs in/near *LAMC2* and *COL25A1* met the genome-wide significance threshold in the discovery dataset, these associations were not observed in the replication dataset, possibly due to the limited statistical power of the second dataset.

Suggestive findings of association in the joint dataset ($p<1E-05$) were seen with SNPs near or in the *PTTG1*, *UBE2L3*, *SLC1A7*, and *KIAA1542* loci, and with rs10737562 (no known gene within 100 kb). *PTTG1*, *KIAA1542*, and *UBE2L3* have been shown to be

Table 1. Characteristics of the study participants.

	anti–dsDNA +	anti–dsDNA –	Healthy controls
Characteristic	(n = 811)	(n = 906)	(n = 4813)
Male, n (%)	67 (8)	40 (4)	1300 (27)
SLE characteristics‡			
Median age at diagnosis, years (interquartile range)	30 (22–40)	37 (28–45)	
Malar rash*, n (%)	430 (53)	447 (49)	
Discoid rash*, n (%)	73 (9)	101 (11)	
Photosensitivity*, n (%)	498 (61)	670 (74)	
Oral ulcers*, n (%)	290 (36)	428 (47)	
Arthritis*, n (%)	662 (82)	719 (79)	
Serositis (pericarditis or pleuritis)*, n (%)	327 (40)	346 (38)	
Neurologic (seizure or psychosis)*, n (%)	92 (11)	87 (10)	
Hematologic (leukopenia, lymphopenia, hemolytic anemia, or thrombocytopenia)*, n (%)	547 (68)	489 (54)	
Immunologic (anti–dsDNA, anti-Sm, or antiphospholipid antibodies)*, n (%)	811 (100)	382 (42)	
Renal disorder (proteinuria or cellular casts)*, n (%)	321 (40)	154 (17)	
Anti-nuclear antibody*, n (%)	792 (98)	837 (92)	
Anti-SSA/Ro autoantibody, n (%)#	267 (34)	198 (23)	
Anti-SSB/La autoantibody, n (%)#	122 (16)	91 (10)	
Anti-Sm autoantibody, n (%)#	138 (18)	53 (6)	
Anti-RNP autoantibody, n (%)#	192 (25)	112 (13)	

‡ SLE denotes systemic lupus erythematosus.
* A person needs to have 4 out of these 11 critieria to be classified as having SLE [5,6].
# Percentages may not be based on the total number in each anti–dsDNA subgroup due to missing data.
doi:10.1371/journal.pgen.1001323.t001

associated with SLE [7,9], and thus, are likely true associations that are specific for anti–dsDNA + SLE. The associations with *SLC1A7* and rs10737562 have not been previously reported with SLE or anti–dsDNA autoantibody production, and should be replicated in another collection of anti–dsDNA + SLE cases.

Anti–dsDNA + SLE cases (n = 811) were also compared to the combined group of anti–dsDNA – SLE cases and healthy controls (n = 5719) to conduct an analysis maximally powered to identify SNPs only associated with anti–dsDNA + SLE. No new loci (i.e., loci not presented in Table 2) displayed significant or suggestive evidence of association. Also, analyses comparing anti–dsDNA + SLE cases to healthy controls excluding ANA negative subjects showed results similar to Table 2 (data not shown).

Anti–dsDNA negative SLE versus healthy controls

Next, we compared anti–dsDNA – SLE cases to healthy controls in the discovery, replication, and combined "joint analysis" datasets using additive logistic regression models adjusted for population stratification as described previously. Table S1 presents the genomic control inflation factor (λ_{GC}) for each analysis prior to and after adjustment for population stratification. The quantile-quantile and Manhattan plots for the joint analysis are displayed in Figure S2.

Far fewer statistically significant genetic associations were observed. Excluding the MHC, one statistically significant association was observed in the joint analysis of genotyped SNPs when none would have been expected under the null hypothesis ($p<5E-07$). The most significant associations were again seen in the MHC, with rs2301271 (~9 kb downstream from *HLA-DQA2)* as the most significantly associated MHC SNP (OR_{joint} 1.47, 95% CI 1.32–1.63, p = 2.0E-12). In the discovery dataset, no SNPs outside of the MHC met our genome-wide significance threshold. In the joint analysis (Table 3), an additional association with rs10488631 near *IRF5* met genome-wide significance (OR_{joint} 1.57, 95% CI 1.35–1.82, p = 6.2E-09). Three SNPs had suggestive evidence of association in the joint analysis: rs2669010 in *RPL7AP50*, rs918959 (no known gene within 100 KB), and the missense SNP rs7927370 in *OR4A15*. These novel findings need to be replicated in another collection of anti–dsDNA – SLE cases. Analyses excluding ANA negative subjects showed similar results (data not shown).

Case-only analysis

Using the combined dataset, we compared the anti–dsDNA + SLE cases (n = 811) to the anti–dsDNA – SLE cases (n = 906) using additive logistic regression models. Minimal population stratification was observed between these two groups ($\lambda_{GC} = 1.01$) without adjustment using principal components. However, we included the principal components in these models to decrease the possible influence of subtle stratification on our findings ($\lambda_{GC} = 1.00$ after adjustment for population stratification).

No SNP met our genome-wide significance threshold ($p<5E-07$) for anti–dsDNA autoantibody production in this analysis. Six SNPs showed suggestive evidence of association ($p<1E-05$), as shown in Table 4. Only three SNPs would be expected to have a $p<1E-05$ under the null hypothesis. Similar to the anti–dsDNA + analysis described above, rs7574865 in *STAT4* was once again found to be associated with anti–dsDNA + SLE. rs1463525 in *NAALADL2* is of interest, since this gene was recently identified as a susceptibility locus for Kawasaki disease [17], another autoim-

Table 2. Loci with significant (p<5E-07) or suggestive (p between 5E-07 and 1E-05) evidence for association with anti–dsDNA + SLE identified in the joint analysis.

					Discovery dataset		Replication dataset		Joint analysis			
					605 cases, 3334 controls		206 cases, 1479 controls		811 cases, 4813 controls			
Gene/ Region	chr	SNP	position	MA‡	OR (95% CI)	p (λ_{GC} = 1.045)	OR (95% CI)	p (λ_{GC} = 1.038)	MAF‡ cases	MAF‡ controls	OR (95% CI)	p (λ_{GC} = 1.048)
TNXB[a]	6	rs1150754	32158736	A	2.05 (1.74–2.40)	9.1E-18	2.77 (2.14–3.58)[b]	1.4E-13[b]	0.25	0.13	2.21 (1.93–2.53)[c]	6.4E-29[c]
HLA-DR3[a]	6	rs2187668	32713862	A	2.14 (1.81–2.52)	3.1E-18	2.54 (1.95–3.32)	1.8E-11	0.22	0.12	2.23 (1.94–2.57)	5.8E-28
STAT4	2	rs7574865	191790139	T	1.74 (1.51–1.99)	2.3E-14	1.87 (1.49–2.36)	1.7E-07	0.35	0.23	1.77 (1.57–1.99)	2.0E-20
IRF5	7	rs10488631	128188134	C	1.73 (1.45–2.05)	1.2E-09	2.52 (1.91–3.32)	1.5E-10	0.19	0.11	1.92 (1.66–2.22)	7.2E-18
ITGAM	16	rs9888739	31220754	T	1.57 (1.33–1.86)	2.8E-07	2.50 (1.93–3.24)	1.1E-11	0.21	0.13	1.80 (1.56–2.07)	1.1E-15
BLK	8	rs2736340	11381382	T	1.30 (1.13–1.50)	3.7E-04	1.59 (1.26–1.99)	9.5E-05	0.30	0.23	1.38 (1.23–1.56)	2.5E-07
LAMC2	1	rs525410	179908087	G	0.71 (0.62–0.8)	2.5E-07	0.89 (0.72–1.1)	0.30	0.45	0.51	0.75 (0.68–0.84)	6.5E-07
COL25A1	4	rs4956211	110080730	A	1.41 (1.24–1.6)	4.2E-07	1.14 (0.91–1.42)[b]	0.27[b]	0.42	0.35	1.33 (1.19–1.48)[c]	1.1E-06[c]
PTTG1	5	rs2431697	159812556	C	0.73 (0.64–0.84)	6.8E-06	0.82 (0.66–1.01)	0.069	0.37	0.44	0.76 (0.68–0.85)	1.7E-06
UBE2L3	22	rs5754217	20264229	T	1.37 (1.17–1.60)	9.7E-05	1.45 (1.14–1.84)	0.0031	0.24	0.19	1.38 (1.21–1.57)	1.9E-06
SLC1A7	1	rs6695567	53341606	A	0.80 (0.70–0.91)	0.0010	0.70 (0.56–0.87)	0.0015	0.37	0.43	0.76 (0.68–0.85)	3.6E-06
KIAA1542	11	rs4963128	579564	A	0.78 (0.68–0.90)	6.5E-04	0.66 (0.52–0.84)	7.5E-04	0.27	0.34	0.75 (0.67–0.85)	3.9E-06
[d]	1	rs10737562	186455280	A	1.02 (0.90–1.16)	0.77	0.36 (0.28–0.45)[b]	4.1E-15[b]	0.44	0.50	0.77 (0.69–0.86)[c]	6.8E-06[c]

Only the most significantly associated SNP per locus is displayed. Additional associated SNPs for each locus are presented in Table S2.
‡ MA = Minor allele, MAF = minor allele frequency.
a Only the most significantly associated MHC SNP and the HLA-DR3 tagSNP is included.
b imputed SNP, λ_{GC} = 1.089.
c imputed SNP, λ_{GC} = 1.071.
d No known gene within 100 kb.
doi:10.1371/journal.pgen.1001323.t002

mune disease. However, the most significantly associated SNP for Kawasaki disease (rs17531088) is not in linkage disequilibrium with the SNP identified in our analysis (r^2 = 0.002 in the CEU HapMap population). The statistical power for this analysis was limited by our relatively smaller sample size. Thus, additional studies are needed to fully explore this area and to replicate our findings.

Since two of the suggested SNPs are located within the MHC, our findings indicate that MHC associations may be heterogeneous between these two subgroups of SLE. This finding is further supported by a plot of the p-values for association among the MHC SNPs, as shown in Figure 2. The strongest MHC associations with anti–dsDNA autoantibody production among these SLE patients were within the class II region. Given the extensive linkage disequilibrium of the MHC, many of these associations may be driven by the MHC class II locus *HLA-DRB1*, a well established SLE susceptibility gene [18,19].

Comparison of SLE–associated SNPs

Next, we examined the magnitude of association between 22 polymorphisms previously associated with SLE in Gateva et al. [7], stratified by anti–dsDNA autoantibody status using tests of heterogeneity. For each SNP, the association result in the anti–dsDNA + versus healthy control analysis was compared to the association result for the anti–dsDNA – versus healthy control analysis. A p-value of less than 0.05 was considered significant evidence of heterogeneity.

Table 5 presents the results of the tests of heterogeneity, along with the association results from the case-only analysis, for these 22 SLE susceptibility loci. Associations for HLA-DR3 (indicated by its tagSNP rs2187668) and SNPs in *STAT4* and *ITGAM* differed substantially between the two anti–dsDNA subgroups ($p_{heterogeneity}$<0.005). In addition, SNPs in the *BANK1*, *KIAA1542*, *ITGAM*, and *UBE2L3* regions also showed differential associations in the two anti–dsDNA subgroups (p<0.05). For all of these SNPs, the associations with anti–dsDNA + SLE had stronger OR and smaller p-values when compared to anti–dsDNA – SLE or SLE itself. The differences are best demonstrated by rs7574865 in *STAT4*: OR for anti–dsDNA + SLE 1.77 (95% CI 1.57–1.99, p = 2.0E-20) compared to OR for anti–dsDNA – SLE 1.26 (95% 1.12–1.41, p = 2.4E-4), with p-value for the test of heterogeneity <0.0005. In contrast, ORs of association were quite similar

Table 3. SNPs with significant (p<5E-07) or suggestive (p between 5E-07 and 1E-05) evidence for association with anti–dsDNA – SLE identified in the joint analysis.

					Discovery dataset (673 cases, 3334 controls)		Replication dataset (233 cases, 1479 controls)		Joint analysis (906 cases, 4813 controls)			
Gene/ Region	chr	SNP	position	MA‡	OR (95% CI)	p (λ_{GC} = 1.029)	OR (95% CI)	p (λ_{GC} = 1.049)	MAF ‡cases	MAF ‡ controls	OR (95% CI)	p (λ_{GC} = 1.045)
HLA-DQA2[a]	6	rs2301271	32833171	T	1.48 (1.3–1.67)	1.4E-09	1.5 (1.23–1.83)	1.2E-04	0.51	0.40	1.47 (1.32–1.63)	2.0E-12
IRF5	7	rs12531711	128211417	G	1.45 (1.21–1.73)	5.4E-05	1.95 (1.48–2.58)[b]	7.2E-06[b]	0.16	0.11	1.58 (1.36–1.83)[c]	6.0E-09[c]
IRF5	7	rs10488631	128188134	C	1.46 (1.22–1.74)	3.8E-05	1.88 (1.42–2.49)	1.5E-05	0.16	0.11	1.57 (1.35–1.82)	6.2E-09
IRF5	7	rs12537284	128311857	A	1.33 (1.13–1.57)	8.8E-04	1.65 (1.27–2.13)	2.1E-04	0.18	0.13	1.42 (1.23–1.62)	1.5E-06
[d]	2	rs918959	181339235	A	0.53 (0.41–0.69)	3.7E-06	0.68 (0.45–1.02)[b]	0.076[b]	0.05	0.09	0.58 (0.46–0.72)[c]	2.4E-06[c]
RPL7AP59	12	rs2669010	75511528	A	1.26 (1.11–1.42)	3.2E-04	1.33 (1.09–1.63)	0.0054	0.45	0.40	1.28 (1.16–1.42)	4.7E-06
OR4A15	11	rs7927370	54892795	T	0.51 (0.36–0.7)	7.1E-05	0.56 (0.33–0.95)	0.034	0.03	0.06	0.52 (0.39–0.69)	7.1E-06

‡ MA = minor allele; MA = minor allele frequency.
a Only the most significantly associated MHC SNP is included.
b imputed analysis, λ_{GC} = 1.097.
c imputed analysis, λ_{GC} = 1.066.
d no known gene within 100 kb.
doi:10.1371/journal.pgen.1001323.t003

between the 2 SLE subgroups for SNPs in/near the *FCGR2A*, *OX40L*, *IL10*, *PXK*, *UHRF1BP1*, *PRDM1*, *BLK*, and *IRAK1* regions.

When examining these SNPs in the case-only analysis, rs2476601 (*PTPN22*), rs10488631 (*IRF5*), and rs2431099 (*PTTG1*) were more strongly associated with anti–dsDNA + SLE than anti–dsDNA – SLE (p<0.05). Sensitivity analysis of the 722 SLE cases with longitudinal anti–dsDNA autoantibody data (of which 46% were anti–dsDNA +, see Methods) showed good consistency in OR with the analyses performed using the full dataset (data not shown).

Among the SNPs studied in this comparison, we did not identify a single SNP that was more strongly associated with anti–dsDNA – SLE disease than anti–dsDNA + SLE or SLE itself, based on OR or p-values.

Genetic Risk Score analysis

To study the relationship between cumulative genetic risk and anti–dsDNA autoantibody production, we calculated an SLE genetic risk score (GRS) by counting the total number of risk alleles an individual had for the 22 SLE-associated SNPs listed in Table 5. The mean SLE GRS was higher in anti–dsDNA + SLE cases (15.5, SD 3.1) compared to anti–dsDNA – SLE cases (14.5, SD 3.0) and healthy controls (13.1, SD 2.8), and the trend was highly statistically significant (p_{trend} = 1.0E-102). In logistic regression analyses adjusting for study source and population stratification, the odds of producing anti–dsDNA among SLE cases increased by 12% (OR 1.12, 95% CI 1.09–1.16) for each 1 unit increase in the SLE GRS. When comparing to healthy controls, the odds of having anti–dsDNA + SLE increased by 32% (OR

Table 4. Loci with suggestive (p between 5E-07 and 1E-05) evidence for association with anti–dsDNA + SLE identified in the case-only analysis (811 anti–dsDNA + SLE cases compared to 906 anti–dsDNA – SLE cases).

					MAF‡	MAF‡		
Gene/Region	chr	SNP	position	MA‡	anti–dsDNA +	anti–dsDNA –	OR (95% CI)	p (λ_{GC} = 1.000)
NOTCH4/C6orf10	6	rs3130320	32331236	T	0.46	0.39	1.39 (1.21–1.60)	3.15E-06
[a]	3	rs12629106	178577774	T	0.13	0.19	0.64 (0.53–0.77)	3.71E-06
NAALADL2	3	rs1463525	176444089	C	0.35	0.43	0.73 (0.63–0.84)	7.69E-06
STAT4	2	rs7574865	191790139	T	0.35	0.28	1.41 (1.21–1.63)	7.95E-06
TMC2	20	rs6049839	2466565	T	0.47	0.40	1.37 (1.19–1.58)	8.16E-06
HLA-DQA1/HLA-DQA2	6	rs2647012	32772436	A	0.51	0.44	1.38 (1.20–1.59)	8.39E-06

‡ MA = minor allele; MAF = minor allele frequency.
[a] no known gene within 100 kb.
doi:10.1371/journal.pgen.1001323.t004

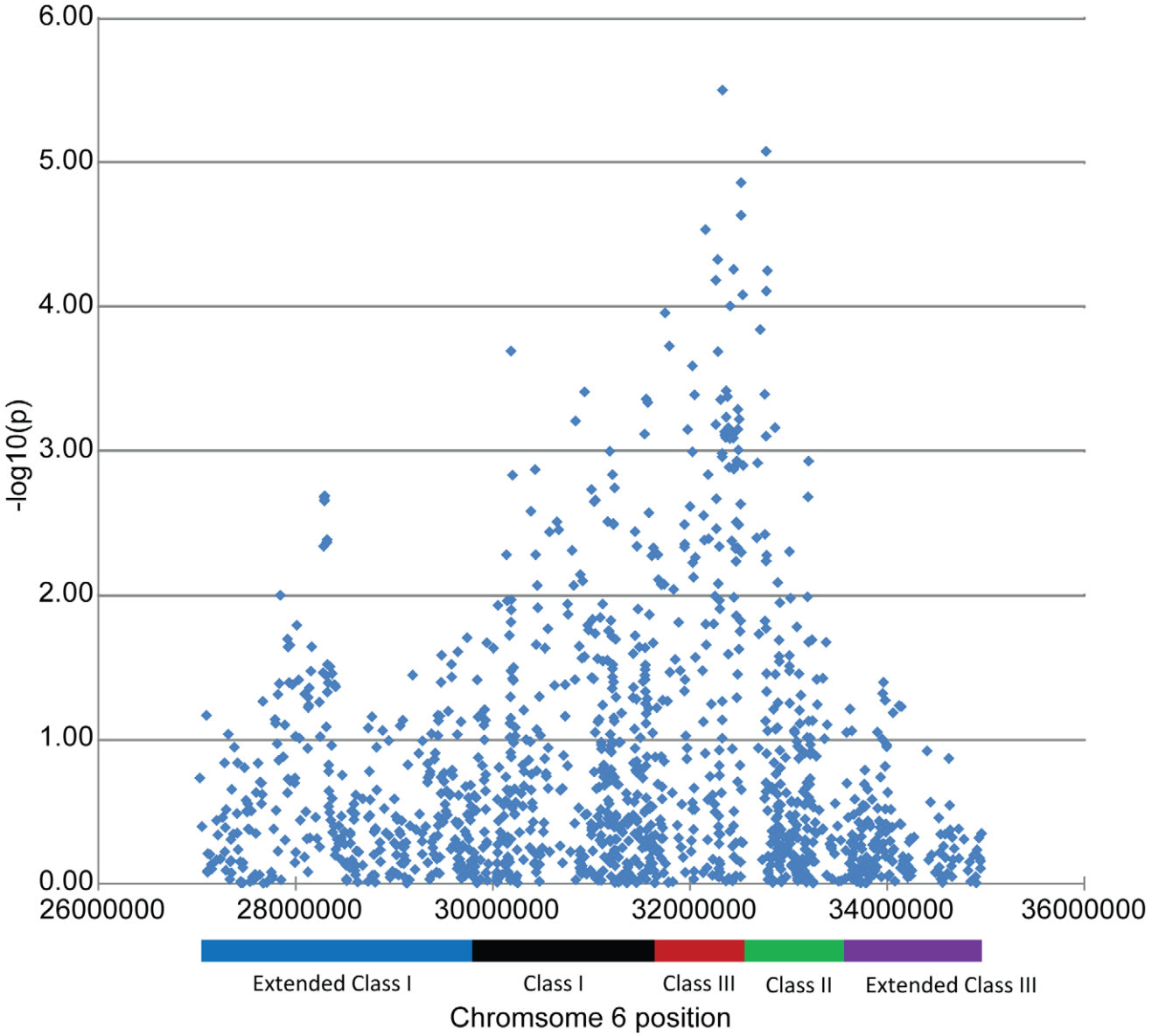

Figure 2. Association results for the MHC SNPs from the case-only analysis.
doi:10.1371/journal.pgen.1001323.g002

1.32, 95% CI 1.28–1.35) for each 1 unit increase in the SLE GRS versus 18% (OR 1.18, 95% CI 1.15–1.21) for anti–dsDNA – SLE. These findings indicate that SLE cases with higher genetic risk are more likely to be anti–dsDNA positive. A more thorough investigation of the association between SLE GRS and SLE manifestations is presented in Taylor et al. [20].

Discussion

In this paper, we present the first GWAS of anti–dsDNA autoantibody production in SLE. We have shown that SNPs in the MHC, *STAT4*, *IRF5*, and *ITGAM* regions are associated with anti–dsDNA + SLE. Only SNPs in the MHC and *IRF5* met genome-wide significance threshold levels in the analysis of anti–dsDNA – SLE, with lower OR and larger p-values compared to their associations with anti–dsDNA + SLE. Furthermore, many of the previously identified SLE susceptibility loci showed differential associations between the 2 anti–dsDNA subgroups. Using a genetic risk score analysis, we found that SLE cases with a greater number of risk alleles were more likely to be anti–dsDNA +. These results suggest that genetic factors may have a greater influence in the development of anti–dsDNA + SLE as compared to anti–dsDNA – SLE.

The strongest association signals for both the anti–dsDNA + and anti–dsDNA – analyses were observed with MHC SNPs. Previous studies have shown that the strongest, most consistent genetic signals with SLE have been with the HLA-DR2 and HLA-DR3 MHC serotypes [18,19]. While we confirm these findings, we also show that the HLA-DR3 association with SLE (as suggested by its tagSNP, rs2187668) is far stronger in anti–dsDNA + SLE as compared to anti–dsDNA – SLE or SLE itself. Thus, the HLA-DR3 allele may have a greater impact on the propensity to produce autoantibodies compared to SLE susceptibility generally. Although a similar finding was observed with HLA-DR2 (tagSNP rs9271366), the test of heterogeneity was not statistically significant, possibly due in part to decreased statistical power

Table 5. Comparison of association results for loci associated with systemic lupus erythematosus described in Gateva et. al [7], stratified by anti–dsDNA autoantibody production.

		SLE versus healthy controls		anti–dsDNA + versus healthy controls		anti–dsDNA – versus healthy controls			Case Only Analysis	
		(1717 cases, 4813 controls)		(811 cases, 4813 controls)		(906 cases, 4813 controls)		Test of heterogeneity	(811 anti–dsDNA+ cases versus 906 anti–dsDNA – cases)	
Gene	SNP	OR (95% CI)	p (λ_{GC} = 1.080)	OR (95% CI)	p (λ_{GC} = 1.048)	OR (95% CI)	p (λ_{GC} = 1.045)	p	OR (95% CI)	p (λ_{GC} = 1.000)
PTPN22	rs2476601	1.26 (1.1–1.43)	9.5E-04	1.41 (1.19–1.66)	8.7E-05	1.14 (0.96–1.35)	0.15	0.078	1.24 (1.00–1.53)	0.047
FCGR2A	rs1801274	0.84 (0.77–0.91)	4.7E-05	0.83 (0.74–0.92)	0.0010	0.84 (0.76–0.94)	0.0022	0.80	0.99 (0.87–1.14)	0.92
OX40L	rs2205960	1.24 (1.13–1.36)	1.7E-05	1.28 (1.13–1.44)	1.6E-04	1.22 (1.09–1.38)	0.0012	0.64	1.06 (0.91–1.23)	0.48
IL10	rs3024505	1.25 (1.13–1.39)	4.8E-05	1.29 (1.12–1.49)	4.1E-04	1.24 (1.08–1.41)	0.0022	0.66	1.04 (0.87–1.23)	0.69
STAT4	rs7574865	1.49 (1.36–1.63)	3.6E-16	1.77 (1.57–1.99)	2.0E-20	1.26 (1.12–1.41)	2.4E-04	<0.0005	1.41 (1.21–1.63)	7.9E-06
PXK	rs6445975	1.14 (1.05–1.25)	0.0045	1.15 (1.02–1.29)	0.027	1.14 (1.02–1.28)	0.027	0.95	1.01 (0.88–1.17)	0.87
BANK1	rs10516487	0.93 (0.85–1.01)	0.11	0.85 (0.76–0.96)	0.011	1.01 (0.9–1.13)	0.92	0.048	0.85 (0.73–0.99)	0.030
TNIP1	rs10036748	1.26 (1.15–1.38)[a]	2.5E-06[a]	1.30 (1.15–1.46)[b]	3.0E-05[b]	1.23 (1.09–1.38)[c]	7.9E-04[c]	0.51	1.08 (0.93–1.25)[d]	0.32[d]
PTTG1	rs2431099	0.84 (0.78–0.91)[a]	1.1E-04[a]	0.78 (0.70–0.87)[b]	1.4E-05[b]	0.91 (0.82–1.01)[c]	0.072[c]	0.052	0.87 (0.76–0.99)[d]	0.037[d]
HLA-DR2	rs9271366	1.27 (1.13–1.41)	7.4E-05	1.36 (1.17–1.57)	7.1E-05	1.18 (1.03–1.36)	0.024	0.19	1.16 (0.97–1.39)	0.11
HLA-DR3	rs2187668	1.91 (1.71–2.14)	1.0E-27	2.23 (1.94–2.57)	5.8E-28	1.62 (1.4–1.87)	1.2E-10	0.002	1.40 (1.18–1.67)	1.4E-04
UHRF1BP1	rs9462015	1.27 (1.16–1.38)	1.8E-07	1.27 (1.14–1.42)	5.3E-05	1.27 (1.14–1.42)	2.4E-05	1.00	1 (0.87–1.15)	0.96
PRDM1	rs6568431	1.19 (1.1–1.3)	4.2E-05	1.19 (1.06–1.32)	0.0026	1.2 (1.08–1.33)	8.5E-04	0.88	0.98 (0.86–1.12)	0.77
ATG5	rs633724	1.13 (1.04–1.22)	0.0069	1.18 (1.06–1.32)	0.0043	1.08 (0.97–1.2)	0.16	0.27	1.1 (0.96–1.26)	0.19
TNFAIP3	rs2327832	1.2 (1.08–1.32)	5.4E-04	1.27 (1.12–1.44)	2.8E-04	1.13 (1–1.28)	0.065	0.19	1.12 (0.96–1.31)	0.14
JAZF1	rs1635852	1.15 (1.06–1.24)	0.0012	1.21 (1.09–1.35)	6.2E-04	1.09 (0.99–1.21)	0.091	0.18	0.89 (0.78–1.02)	0.091
IRF5	rs10488631	1.74 (1.55–1.95)	1.9E-19	1.92 (1.66–2.22)	7.2E-18	1.57 (1.35–1.82)	6.2E-09	0.056	1.23 (1.03–1.48)	0.022
BLK	rs2248932	1.24 (1.14–1.35)	2.7E-06	1.28 (1.14–1.43)	2.9E-05	1.2 (1.08–1.34)	0.0011	0.44	1.06 (0.92–1.22)	0.39
KIAA1542	rs4963128	0.84 (0.77–0.92)	1.9E-04	0.75 (0.67–0.85)	3.9E-06	0.92 (0.82–1.03)	0.14	0.014	0.81 (0.70–0.93)	0.0034
ITGAM	rs9888739	1.53 (1.37–1.71)	3.4E-13	1.8 (1.56–2.07)	1.1E-15	1.32 (1.15–1.53)	2.1E-04	0.003	1.35 (1.14–1.61)	7.0E-04
UBE2L3	rs5754217	1.22 (1.1–1.35)	1.7E-04	1.38 (1.21–1.57)	1.9E-06	1.09 (0.95–1.24)	0.22	0.011	1.28 (1.08–1.51)	0.0038
IRAK1	rs2269368	1.19 (1.06–1.33)	6.1E-03	1.2 (1.03–1.4)	2.2E-02	1.17 (1.01–1.35)	0.045	0.80	1.02 (0.85–1.23)	0.82

a imputed analysis, λ_{GC} = 1.118.
b imputed analysis, λ_{GC} = 1.071.
c imputed analysis, λ_{GC} = 1.066.
d imputed analysis, λ_{GC} = 1.001.
doi:10.1371/journal.pgen.1001323.t005

since the DR2 tagSNP is less common than the DR3 tagSNP (DR2 tagSNP minor allele frequency 0.182 in anti–dsDNA + SLE, 0.167 in anti–dsDNA – SLE, and 0.143 in healthy controls). Examination of other MHC SNPs in the case-only analysis indicates that other (non-*HLA-DRB1*) loci may have associations with anti–dsDNA autoantibody production beyond the associations observed with SLE.

In addition to the HLA-DR3 tagSNP discussed above, the associations between the *STAT4* and *ITGAM* SNPs and anti–dsDNA + SLE were stronger in magnitude than the associations with SLE per se in our datasets (Table 5). The smaller p-values seen in the associations for these loci with anti–dsDNA + SLE are especially striking given the substantially smaller sample size of this subgroup. Our results imply that *STAT4*, *ITGAM*, and HLA-DR3 may be more accurately considered "autoantibody propensity loci" rather than simply "SLE susceptibility loci" given their significant tests of heterogeneity ($p<0.05$). Using this criterion, three other previously identified SLE susceptibility loci may also be considered autoantibody propensity loci: *KIAA1542*, *BANK1*, and *UBE2L3*. In fact, these SNPs had no evidence of association with anti–dsDNA – SLE in this study ($p>0.05$). By characterizing these SNPs as autoantibody propensity loci, we identify a potential mechanistic role for these disease associations.

Are these autoantibody propensity loci associated with other autoantibodies? In rheumatoid arthritis (RA), other alleles of the *HLA-DRB1* locus (collectively referred to as the "shared epitope") are associated with anti-CCP autoantibody positivity [21]. While a study of *STAT4* (rs7574865) in an early RA inception cohort suggested an association with the anti-CCP autoantibody [22], others have not a shown strong association between this SNP and seropositivity in RA [23]. *PTPN22* (rs2476601) has been shown to be more strongly associated with autoantibody positive RA [22]. In our study, other SLE-related autoantibodies (anti-SSA, anti-SSB, anti-Sm, and anti-RNP) are more frequent in the anti–dsDNA + subgroup (Table 1), but correlations between anti–dsDNA and these other autoantibodies antibodies are modest, with Pearson correlation coefficients <0.2 (data not shown). Thus, additional studies are needed to further investigate whether these or other loci are associated with other autoantibodies.

Of note, not all of the previously identified SLE susceptibility SNPs showed differential associations between the anti–dsDNA subgroups. In fact, the OR for the SNPs in or near *FCGR2A*, *OX40L*, *PXK*, and *UHRF1BP1* were strikingly similar between the anti–dsDNA + and anti–dsDNA – subgroups. These loci may represent more generalized SLE susceptibility loci, and their mode of conferring SLE disease risk is likely independent of anti–dsDNA autoantibody production. While *PTPN22* (rs2476601), *IRF5* (rs10488631), and *PTTG1* (rs2431099) do not fulfill our criterion as autoantibody susceptibility loci, the results of the case-only analysis suggest that these loci may have a stronger effect in anti–dsDNA + SLE.

Interestingly, far fewer associations were seen in the anti–dsDNA – SLE analysis. Even in the joint analysis, which had the most statistical power, only 1 SNP outside of the MHC met our genome-wide significance threshold—rs10488631 in *IRF5*. This finding may be explained by a number of different reasons. SNP associations for anti–dsDNA – subgroup may be weaker, and thus would require a larger sample of anti–dsDNA – SLE cases in order to be identified. Other types of genetic variation or non-genetic factors, such as environmental exposures [24,25], may have a stronger influence on susceptibility to anti–dsDNA – SLE. Lastly, the anti–dsDNA – subgroup may be more clinically heterogeneous or be comprised of individuals who develop SLE through different pathogenic (and genetic) mechanisms, thus decreasing our statistical power to identify genetic associations with this subgroup.

One limitation of this study is the potential misclassification of anti–dsDNA autoantibody status. This misclassification may have occurred because the anti–dsDNA autoantibody was assessed by different assays between the participating case collections, and a patient's anti–dsDNA status can vary over the disease course. However, this misclassification would bias our findings of differences between anti–dsDNA + and anti–dsDNA – SLE towards the null. Moreover, sensitivity analyses performed using the available longitudinal data showed consistent ORs, suggesting that the potential misclassification did not greatly influence our results. A second limitation is that all participants were of European descent. Limiting this study to those of European descent minimizes confounding due to genetic differences arising from differences in ethnicity. Future efforts should study non-European populations given their increased incidence of SLE [26,27].

In summary, this GWAS of anti–dsDNA autoantibody production in SLE shows that there are more, and stronger, genetic associations in anti–dsDNA + SLE compared to anti–dsDNA – SLE. Previously identified SLE susceptibility loci such as *STAT4*, *ITGAM*, *KIAA1542*, *BANK1*, and *UBE2L3* are more strongly associated with anti–dsDNA + SLE and may confer disease risk through their role in autoantibody production. Weaker associations in anti–dsDNA – SLE may suggest that other types of genetic variation or non-genetic factors have a greater impact on disease risk. Lastly, focusing genetic studies on clinical disease characteristics decreases the heterogeneity that could cloud association results and may provide greater insight into pathogenic disease mechanisms.

Methods

Ethics statement

Written informed consent was obtained from all study participants and the institutional review board at each collaborating center approved the study.

Subjects and genotyping

For this study, all SLE cases and healthy controls were of European descent. All SLE cases fulfilled at least 4 ACR classification criteria for SLE [5,6]. Figure 1 presents the final sample sizes in the discovery and replication datasets, and the final number of SNPs advanced to analysis.

Discovery dataset. The discovery dataset utilized genotyping data from the GWAS of SLE published by Hom et al. [8]. All SLE cases (n = 1311) and healthy controls (n = 3340) were genotyped for over 500,000 single nucleotide polymorphisms (SNPs) using the Illumina HumanHap550 BeadChip at the Feinstein Institute of Medical Research (Manhasset, NY). SLE subjects were participants from case collections at the following institutions/consortia: University of California, San Francisco (UCSF, n = 595), Autoimmune Biomarkers Collaborative Network (n = 301), University of Pittsburgh (n = 305), and the Multiple Autoimmune Disease Genetics Consortium (n = 110). Control subjects were obtained from the New York Health Project (NYHP, n = 294), and from iControlDB (n = 3046, http://www.illumina.com/science/icontroldb.ilmn).

Replication dataset. The replication dataset utilized genotyping data from GWAS of SLE published by The International Consortium on the Genetics of Systemic Lupus Erythematosus (SLEGEN) [9]. SLE subjects (n = 769) were obtained from case collections based at the following institutions: University of Minnesota (n = 255), Oklahoma Medical Research Foundation (OMRF, n = 239), UCSF (n = 93), University of California, Los Angeles (n = 85), Uppsala University (n = 62), and the University of Southern California (USC, n = 37). Controls (n = 3620) were provided by the following groups/institutions: iControlDB (n = 3032), NYHP (n = 490), OMRF (n = 90), and USC (n = 8). All SLE cases and 588 healthy controls in this dataset were genotyped for 317,501 SNPs on the Illumina Infinium HumanHap300 BeadChip at the Broad Institute Center for Genotyping and Analysis (Cambridge, MA). Genotypes for the remaining 3032 controls were obtained from iControlDB. Since both datasets utilized controls from the same publicly available databases, substantial overlap in controls was suspected. Therefore, we utilized 1142 healthy controls from the breast cancer study sponsored by the Cancer Genetic Markers of Susceptibility (CGEMS) project (http://cgems.cancer.gov/data/) [10] typed on the Illumina HumanHap500 BeadChip as additional controls for the replication dataset.

Data quality filters. The following data quality filters were applied separately to the discovery, replication, and CGEMS datasets: SNPs were removed from analysis if they had greater than 10% missing genotypes, a minor allele frequency less than 1%, or evidence of deviation from Hardy Weinberg equilibrium in the controls ($p<1\times10^{-4}$). Subjects were removed from analysis if their overall genotyping rate was <90%, were population outliers (more than 6 standard deviations from the mean along any of the first 10 principal components described below) or lacked information regarding anti–dsDNA autoantibody status. Duplicates and first degree relatives were identified using identity-by-state measures calculated in PLINK [12] (http://pngu.mgh.harvard.edu/purcell/plink/), and subjects typed in the discovery dataset were preferentially retained.

Imputation. We imputed the SNPs present on the Illumina HumanHap550 BeadChip (used for the discovery dataset) but absent on the Illumina HumanHap300 BeadChip for the SLEGEN cases and controls. Imputation was performed using IMPUTE version 1 [28] (http://www.stats.ox.ac.uk/~marchini/software/gwas/impute.html) and used CEPH subjects from the International HapMap Project release 21 (www.hapmap.org) as the reference. These SNPs were also imputed in 500 cases and 500 controls randomly selected from the discovery dataset. SNPs were removed from analysis if the imputation confidence score or the

concordance between the imputed genotype and the actual genotype in the randomly selected subgroup was less than 90%. The data quality filters described above were also applied to the replication dataset after merging the imputed and assayed genotypes.

Genetic Risk Score. We calculated a SLE genetic risk score (GRS) based on 22 SNPs (or their proxy, if the listed SNP was not genotyped) with previously established evidence of association described in Gateva et al. [7]. The genetic risk score (GRS) was defined as the sum of the number of risk alleles for these SNPs. If a SNP was protective for SLE (OR <1.0), the major allele was considered the risk allele. Since omitting sporadic missing data would underestimate the number of risk alleles, we utilized the most likely genotype from SNP imputation (using the IMPUTE version 2 –pgs_miss option) for these calculations.

Anti–dsDNA autoantibody status

Anti–dsDNA autoantibody status for all SLE cases was determined by medical record review and/or serologic testing of banked serum. Since anti–dsDNA autoantibody status can fluctuate with disease activity, a SLE case had to have at least one definitively positive laboratory result to be considered anti–dsDNA +. A SLE subject was considered anti–dsDNA – if all laboratory results in the medical record and serologic testing for this autoantibody were negative.

Longitudinal anti–dsDNA autoantibody status (i.e., at least 2 individually documented measurements) was available for a subgroup of SLE cases (n = 722). These data were used for a sensitivity analysis, where anti–dsDNA + was defined as having at least 2 positive anti–dsDNA laboratory results in the longitudinal data, and anti–dsDNA – was defined as having all negative laboratory results for this autoantibody in the longitudinal data.

Statistical analysis

Three GWAS were performed: anti–dsDNA + SLE subjects versus healthy controls, anti–dsDNA – SLE subjects versus healthy controls, and anti–dsDNA + SLE subjects versus anti–dsDNA – SLE subjects (referred to as the case-only analysis). To determine if genetic associations were significantly different between the 2 anti–dsDNA subgroups, tests of heterogeneity were performed for previously identified SLE susceptibility loci. For these loci, the case-only analysis was also repeated using only the longitudinal dataset as a sensitivity analysis. Lastly, a genetic risk score analysis was conducted using logistic regression.

For each GWAS, associations with anti–dsDNA autoantibody status were assessed using additive logistic regression models implemented in PLINK and included the first 5 principal components as co-variates to adjust for population stratification. Principal components were calculated using EIGENSTRAT [29] (http://genepath.med.harvard.edu/~reich/Software.htm). After removal of SNPs in regions with extensive linkage disequilibrium on chromosomes 5 (44–51.5 Mb), 6 (25–33.5 Mb), 8 (8–12 Mb), 11 (45–57 Mb), and 17 (40–43 Mb), the remaining SNPs common to all of the genotyping platforms were used to calculate the principal components. The first 5 principal components were selected based on review of the eigenvalues for the first 10 principal components. A plot of the first two principal components for each individual in the study is shown in Figure S3. P-values for association were adjusted for the genomic control inflation factor (λ_{GC}) observed for each analysis after accounting for population stratification (Table S1). Each GWAS also included calculation of the expected p-value distribution using PLINK to determine the expected number of statistically significant SNPs.

Analyses were conducted separately for the discovery and replication datasets. These datasets were then combined into a "joint dataset" for maximal statistical power. The study source (discovery versus replication dataset) was included as a co-variate in analyses of the joint dataset. For the discovery and joint datasets, a p-value of less than 5E-07 was considered statistically significant, and p-values between 5E-07 and 1E-05 were considered suggestive of association. Statistically significant SNPs in the discovery dataset were examined in the replication dataset, where a p-value of less than 0.005 was considered statistically significant. Analyses of the discovery, replication, and joint datasets first used only the assayed SNPs. These analyses were repeated for the replication and joint datasets to include the imputed SNPs that passed the data quality filters, since the λ_{GC} was expected to differ in analyses using imputed SNPs.

Based on the publication by Gateva et al. [7], 22 SNPs with previously established evidence of association with SLE were analyzed further. For these SNPs (or their proxy, if the listed SNP was not genotyped), the association results for the SNP in two of the GWAS conducted above (anti–dsDNA + or - versus healthy controls) were compared using tests of heterogeneity (STATA 9.0/SE, College Station, TX). A p-value of less than 0.05 was considered significant evidence of heterogeneity. The results for these SNPs were also examined for the case-only analysis (including the sensitivity analysis with longitudinal data), where anti–dsDNA + SLE cases were compared to anti–dsDNA – SLE cases. A p-value of less than 0.05 was considered significant evidence of a differential association between the 2 subgroups. For comparison, the association with SLE was assessed using the logistic regression methods described above.

Associations between the SLE GRS and anti–dsDNA status were calculated using logistic regression models (STATA 9.0/SE, College Station, TX). These models utilized the SLE GRS as a continuous predictor, and adjusted for population stratification (using the first 5 principal components) and study source.

Supporting Information

Figure S1 Association results for the GWAS of anti-dsDNA + SLE cases versus healthy controls (joint analysis of genotyped SNPs). (A) Genomic control inflation factor (λ_{GC})-corrected p-values with significance line at p = 5E-07. (B) Quantile-quantile plot with all SNPs and after removing SNPs in the extended MHC region.
Found at: doi:10.1371/journal.pgen.1001323.s001 (0.24 MB TIF)

Figure S2 Association results for the GWAS of anti-dsDNA – SLE cases versus healthy controls (joint analysis of genotyped SNPs). (A) Genomic control inflation factor (λ_{GC})-corrected p-values with significance line at p = 5E-07, displayed with the same y-axis range as in Figure S1. (B) Quantile-quantile plot with all SNPs and after removing SNPs in the extended MHC region.
Found at: doi:10.1371/journal.pgen.1001323.s002 (0.24 MB TIF)

Figure S3 Plot of the first 2 principal components for each subject included in the study.
Found at: doi:10.1371/journal.pgen.1001323.s003 (0.67 MB TIF)

Table S1 Genomic inflation factors (λ_{GC}) for the analyses of genotyped SNPs prior to and after adjustment for population stratification using principal components.
Found at: doi:10.1371/journal.pgen.1001323.s004 (0.01 MB PDF)

Table S2 All SNPs with significant (p<5E-07) or suggestive (p between 5E-07 and 1E-05) evidence for association with anti-dsDNA + SLE identified in the joint analysis.
Found at: doi:10.1371/journal.pgen.1001323.s005 (0.02 MB PDF)

Acknowledgments

We would like to thank all of the SLE participants and the referring physicians for their contributions to our research.

The International Consortium on the Genetics of Systemic Erythematosus (SLEGEN) is comprised of the following members: Marta E. Alarcón-Riquelme (Oklahoma Medical Research Foundation), Lindsey A. Criswell (University of California, San Francisco), John B. Harley (Cincinnati Children's Hospital Medical Center), Chaim O. Jacob (University of Southern California), Robert P. Kimberly (University of Alabama at Birmingham), Carl D. Langefeld (Wake Forest University Health Sciences), Kathy L. Moser (Oklahoma Medical Research Foundation), Betty P. Tsao (University of California, Los Angeles), and Timothy J. Vyse (imperial College London).

Author Contributions

Conceived and designed the experiments: SAC LAC. Performed the experiments: SAC KET JN. Analyzed the data: SAC KET. Contributed reagents/materials/analysis tools: SAC KET RRG JN ATL WAO COJ MEAR BPT JBH PMG KLM MP FYD MIK SM PKG CDL TWB LAC. Wrote the paper: SAC LAC.

References

1. Pisetsky D (2001) Systemic Lupus Erythematosus: Epidemiology, Pathology, and Pathogenesis. In: Klippel J, ed. Primer of the Rheumatic Diseases. 12 ed. Atlanta: Arthritis Foundation.
2. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, et al. (2003) Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 349: 1526–1533.
3. Buyon J (2008) Systemic Lupus Erythematosus: A. Clinical and Laboratory Features. In: Klippel J, ed. Primer on the Rheumatic Diseases. 13 ed. New YorkNY: Springer. pp 303–318.
4. Isenberg DA, Manson JJ, Ehrenstein MR, Rahman A (2007) Fifty years of anti-ds DNA antibodies: are we approaching journey's end? Rheumatology (Oxford) 46: 1052–1056.
5. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, et al. (1982) The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 25: 1271–1277.
6. Hochberg MC (1997) Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 40: 1725.
7. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, et al. (2009) A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet 41: 1228–1233.
8. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, et al. (2008) Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 358: 900–909.
9. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, et al. (2008) Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 40: 204–210.
10. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39: 870–874.
11. Alarcon GS, McGwin G, Jr., Petri M, Reveille JD, Ramsey-Goldman R, et al. (2002) Baseline characteristics of a multiethnic lupus cohort: PROFILE. Lupus 11: 95–101.
12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.
13. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, et al. (2007) STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 357: 977–986.
14. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, et al. (2006) A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet 38: 550–555.
15. Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, et al. (2008) A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti–dsDNA and shows additive effects with two risk alleles of IRF5. Hum Mol Genet 17: 2868–2876.
16. Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, et al. (2008) Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. PLoS Genet 4: e1000084. doi:10.1371/journal.pgen.1000084.
17. Burgner D, Davila S, Breunis WB, Ng SB, Li Y, et al. (2009) A genome-wide association study identifies novel and functionally related susceptibility Loci for Kawasaki disease. PLoS Genet 5: e1000319. doi:10.1371/journal.pgen.1000319.
18. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, et al. (2008) Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genet 4: e1000024. doi:10.1371/journal.pgen.1000024.
19. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, et al. (2009) High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet 5: e1000696. doi:10.1371/journal.pgen.1000696.
20. Taylor KE, Chung SA, Graham RR, Ortmann WA, Lee AT, et al. (2011) Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes. PLoS Genet 7: e1001311. doi:10.1371/journal.pgen.1001311.
21. van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Huizinga TW, Toes RE, et al. (2006) The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 54: 1117–1121.
22. Morgan AW, Robinson JI, Conaghan PG, Martin SG, Hensor EM, et al. (2010) Evaluation of the rheumatoid arthritis susceptibility loci HLA-DRB1, PTPN22, OLIG3/TNFAIP3, STAT4 and TRAF1/C5 in an inception cohort. Arthritis Res Ther 12: R57.
23. Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, et al. (2007) Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 13: 455–460.
24. Cooper GS, Gilbert KM, Greidinger EL, James JA, Pfau JC, et al. (2008) Recent advances and opportunities in research on lupus: environmental influences and mechanisms of disease. Environ Health Perspect 116: 695–702.
25. Sarzi-Puttini P, Atzeni F, Iaccarino L, Doria A (2005) Environment and systemic lupus erythematosus: an overview. Autoimmunity 38: 465–472.
26. McCarty DJ, Manzi S, Medsger TA, Jr., Ramsey-Goldman R, LaPorte RE, et al. (1995) Incidence of systemic lupus erythematosus. Race and gender differences. Arthritis Rheum 38: 1260–1270.
27. Johnson AE, Gordon C, Palmer RG, Bacon PA (1995) The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. Arthritis Rheum 38: 551–558.
28. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906–913.
29. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904–909.